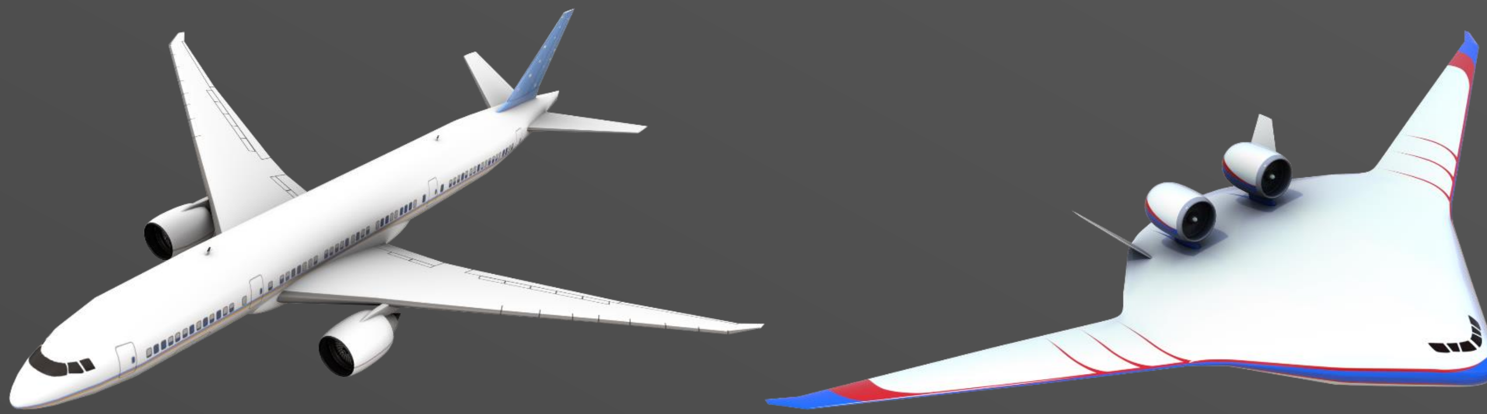
National Aeronautics and Space Administration

Uncertainty estimates of psychoacoustic thresholds obtained from group tests

Jonathan Rathsam
Andrew Christian



Spring 2016 Meeting of the Acoustical Society of America
Salt Lake City, UT
May 24, 2016

Acknowledgments

- NASA Commercial Supersonic Technology Project
  - Jacob Klos, Alexandra Loubeau, Jerry Rouse, Kevin Shepherd
- NASA Environmentally Responsible Aviation Project
  - Steve Rizzi, Russ Thomas, Casey Burley

Outline

1. Research motivation
2. Confidence interval estimation methods
    a. Bayesian Posterior Estimation
    b. Bootstrap (Parametric and Non-Parametric)
    c. Delta Method
3. Results
4. Conclusion

Research motivation

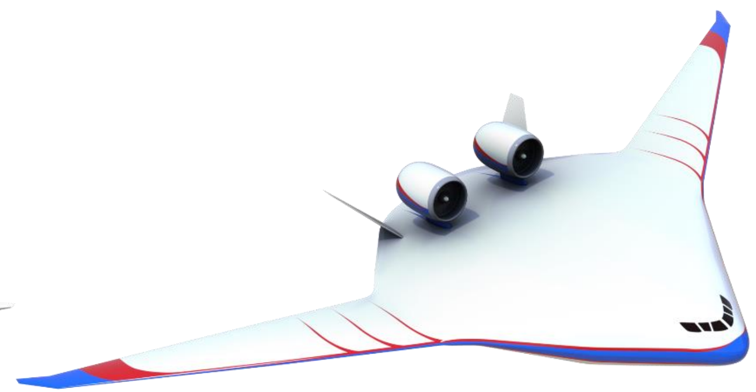
- Find threshold for *binary outcome:*
  - Defaulting on credit card debt (yes/no)
    - based on monthly balance
  - Projectile pierces armor (yes/no)
    - based on projectile velocity
  - Subjects find test signal more annoying than reference signal
    - based on test signal level

- Two research groups with same question

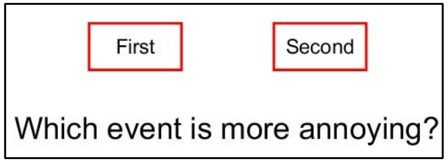
Aircraft Auralizations

Reference

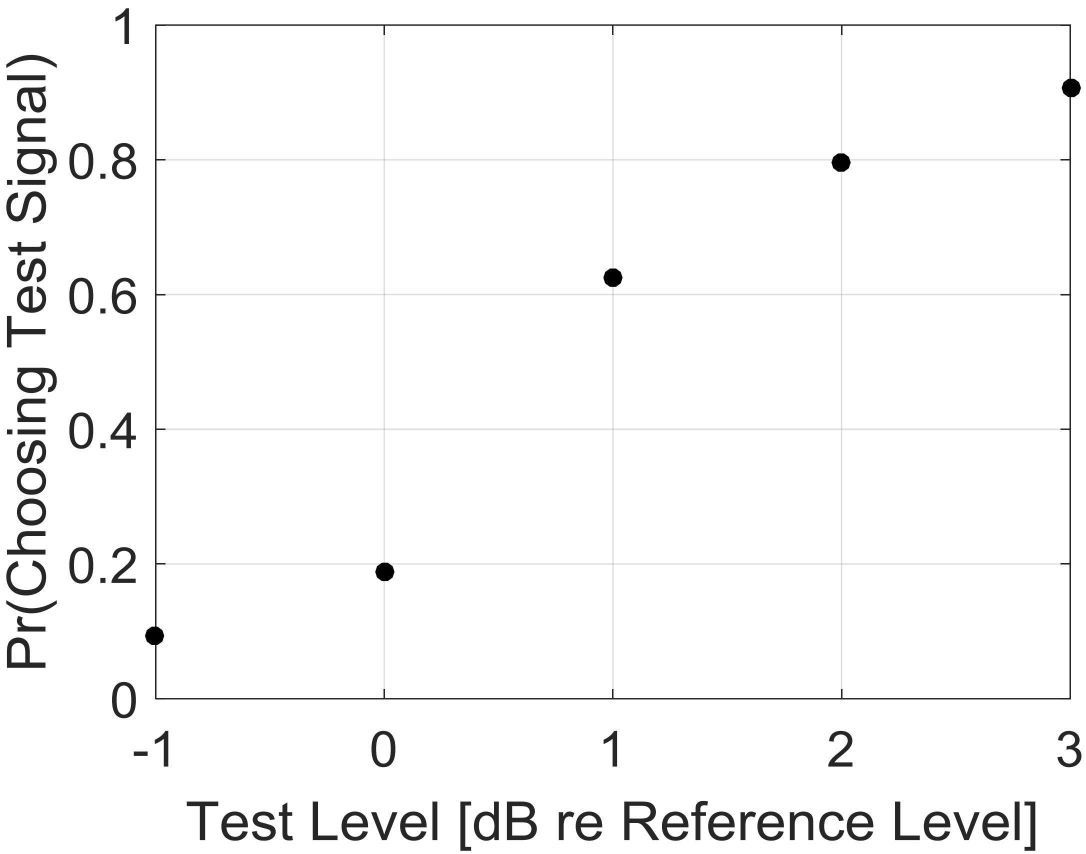
Test



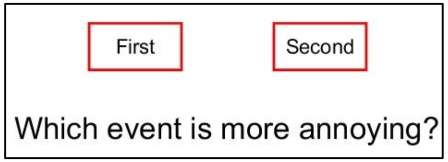
Test Method



First
Second
Which event is more annoying?

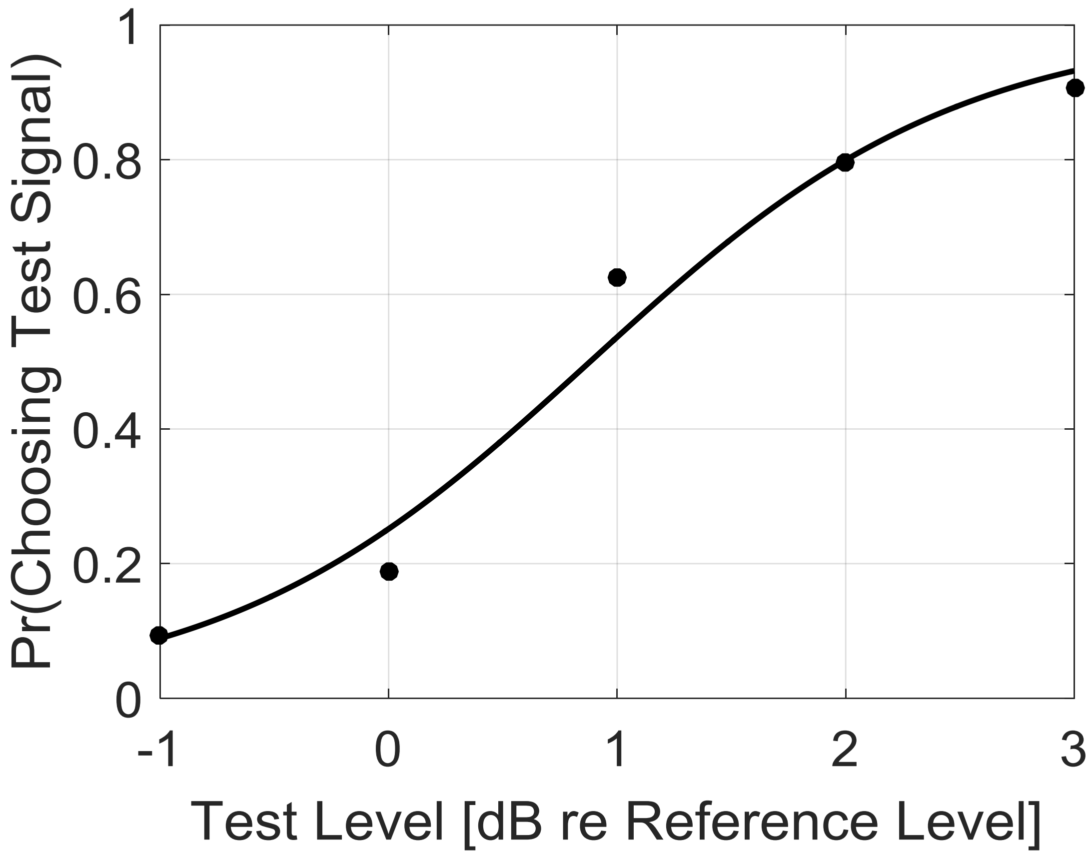


Pr(Choosing Test Signal)
0
0.2
0.4
0.6
0.8
1
-1
0
1
2
3
Test Level [dB re Reference Level]

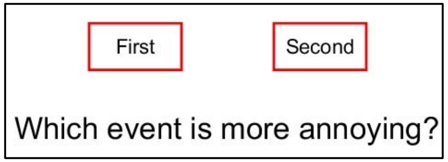


Test Method

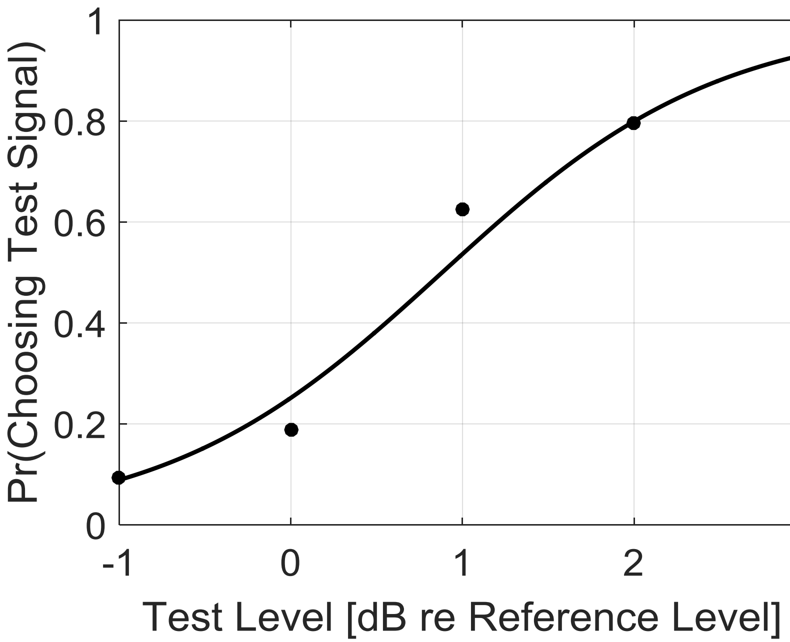
$$\Pr(y_i = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

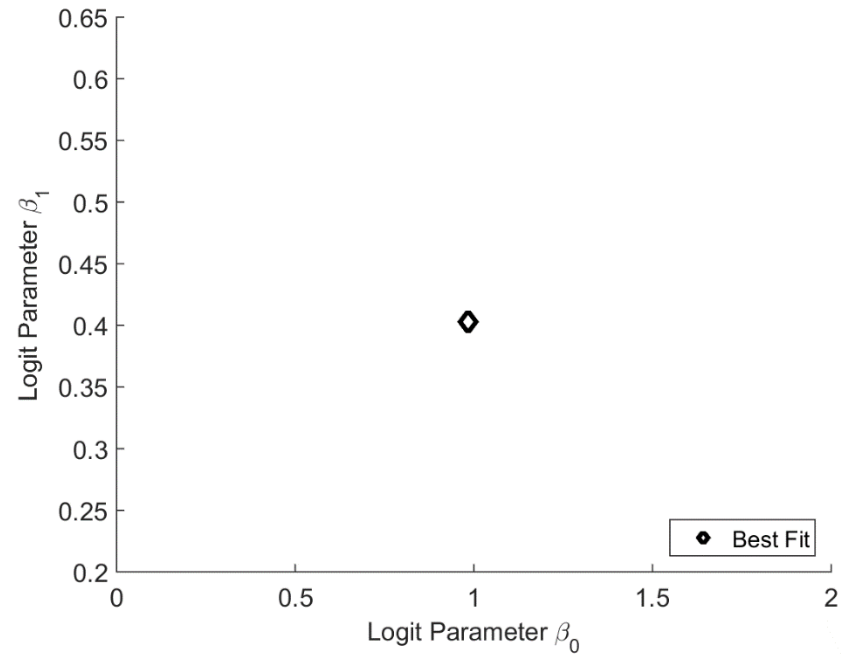


Test Method



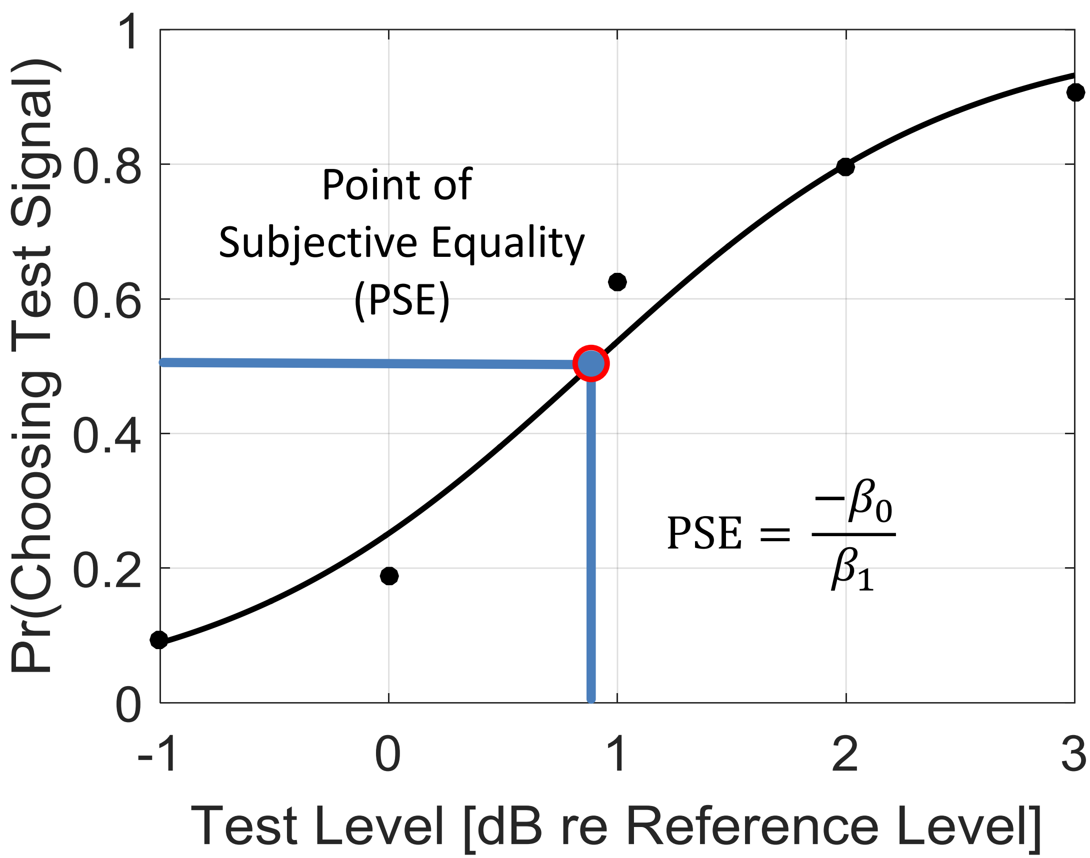
$$\Pr(y_i = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$



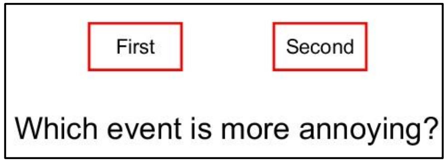


Test Method

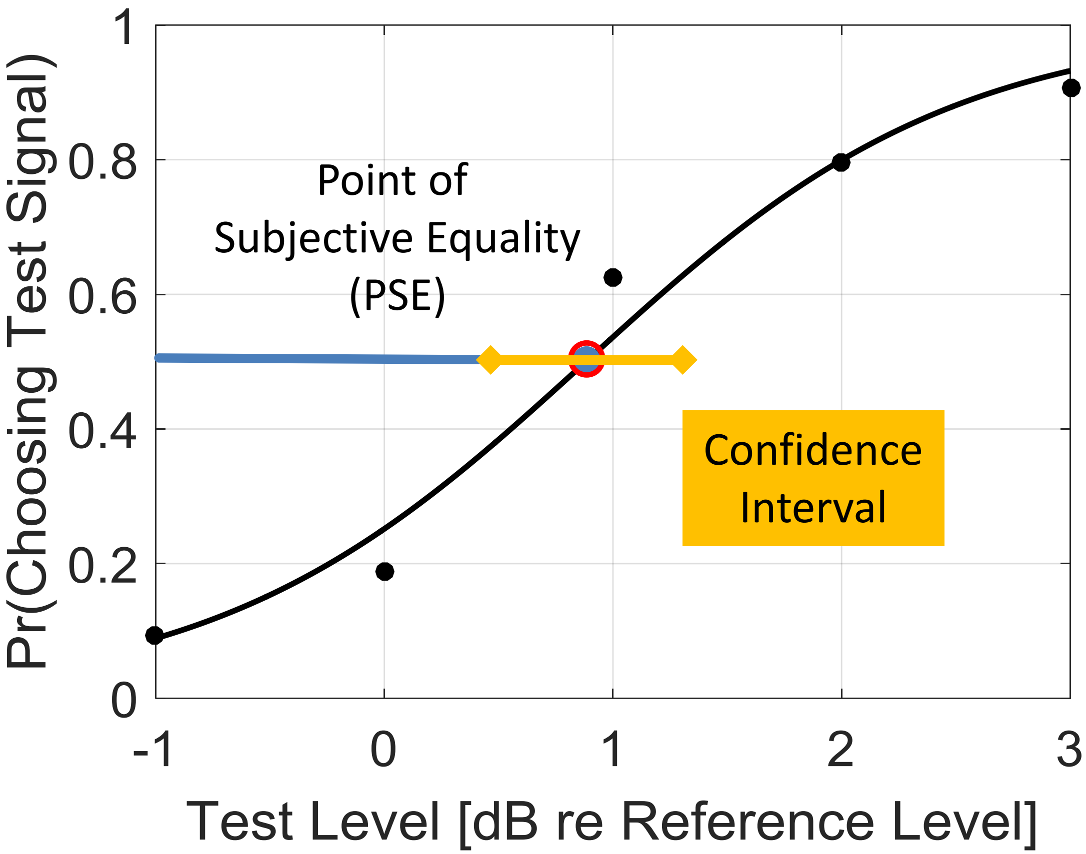
$$\Pr(y_i = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$



Test Method



First
Second
Which event is more annoying?



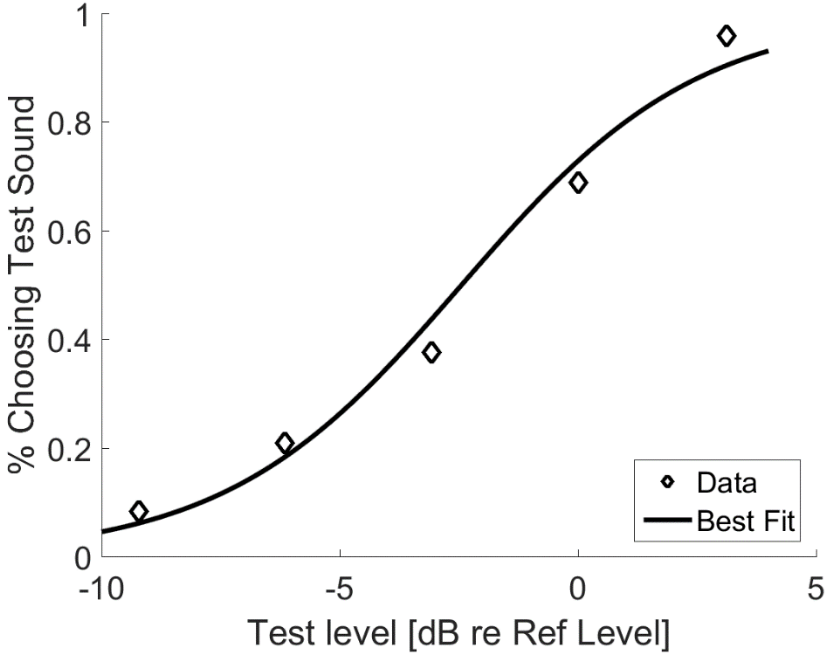
Pr(Choosing Test Signal)
Test Level [dB re Reference Level]
Point of
Subjective Equality
(PSE)
Confidence
Interval
0
0.2
0.4
0.6
0.8
1
-1
0
1
2
3

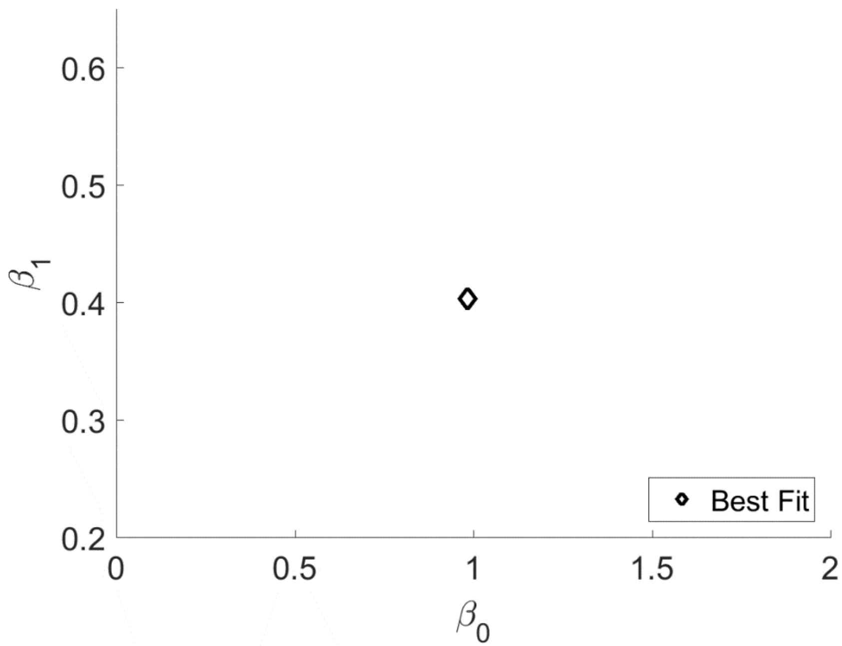
Research Question

- What is most appropriate interval estimation technique?
    a. Bayesian Posterior Estimation
    b. Bootstrap: non-parametric
    c. Bootstrap: parametric
    d. Delta Method

Bayesian Posterior Estimation

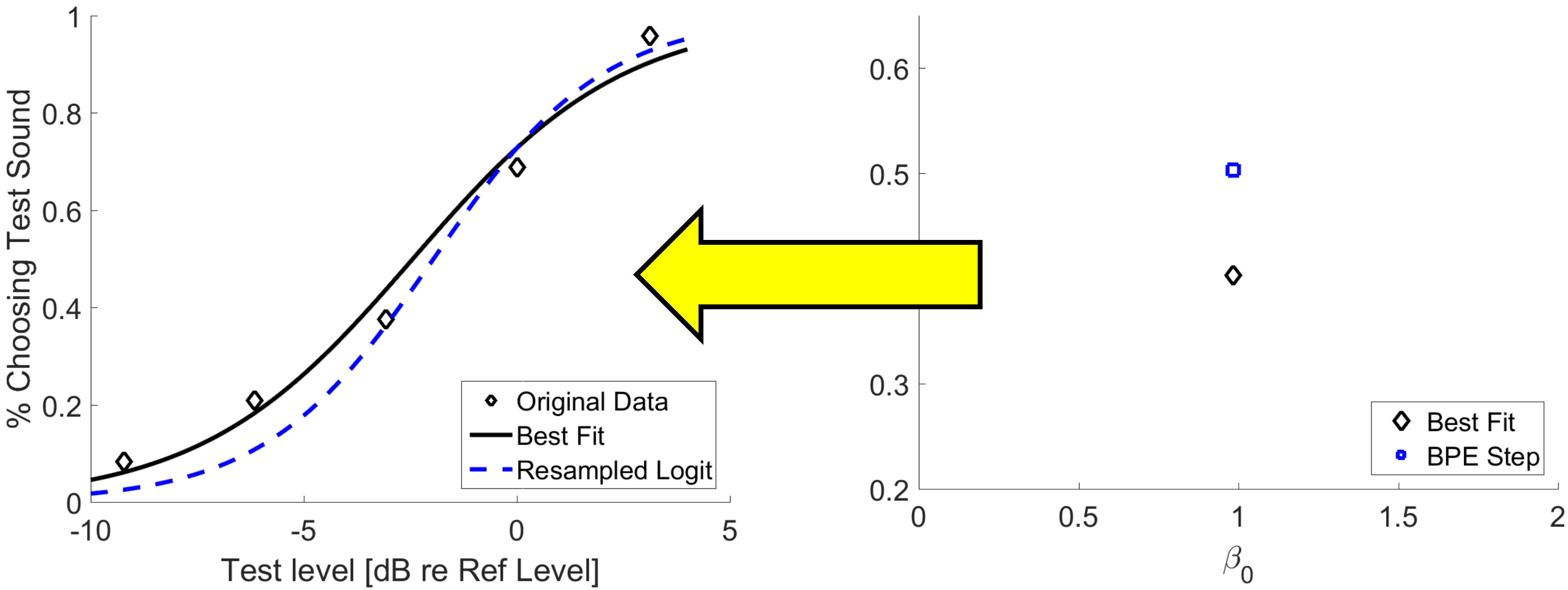
- Begin with data and best fit…



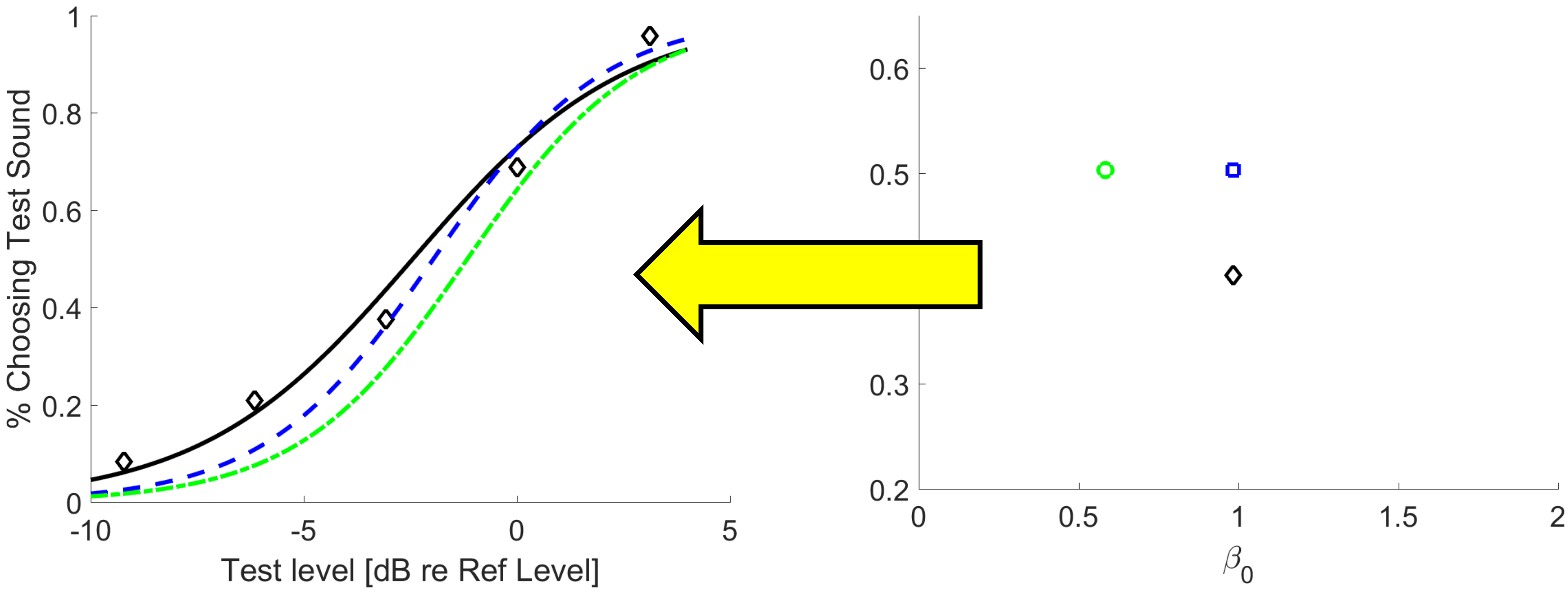


Bayesian Posterior Estimation

Blue line fit is poorer than black line, but still reasonable



Bayesian Posterior Estimation

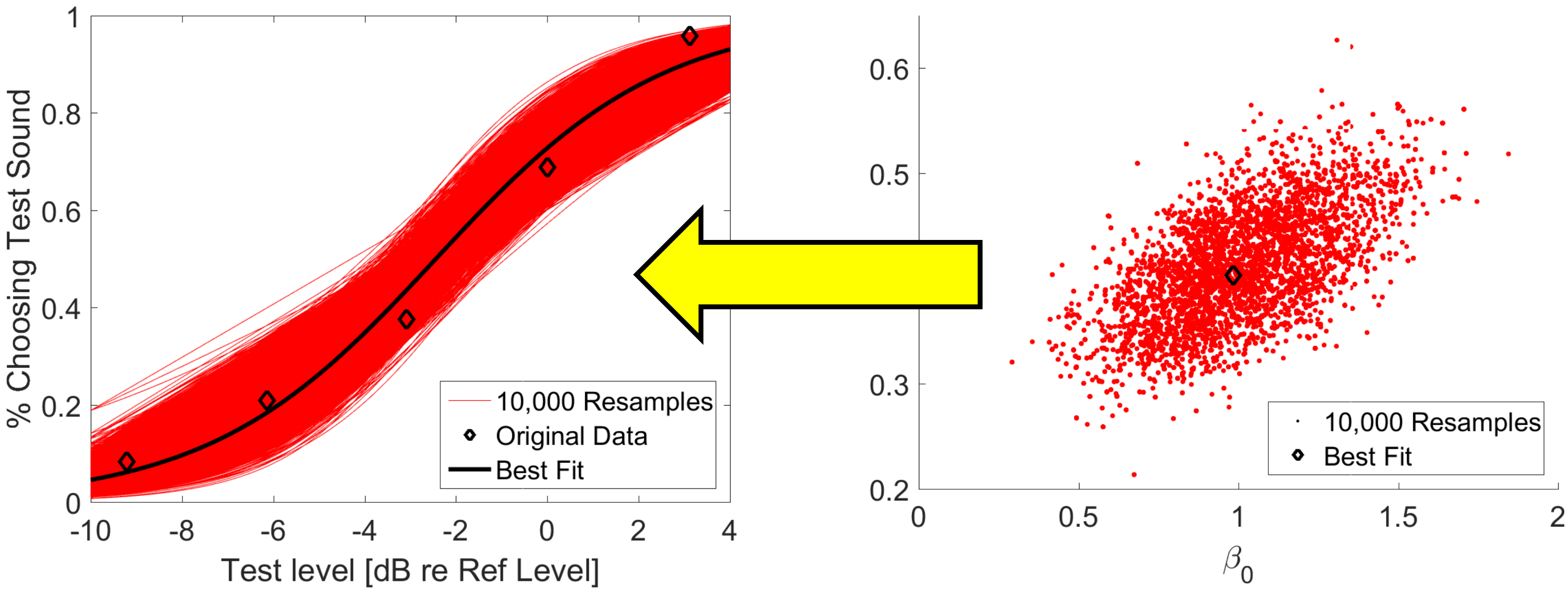


Bayesian Posterior Estimation

BPE numerically samples likelihood function…

allowing confidence intervals to be constructed

anywhere along logistic probability curve

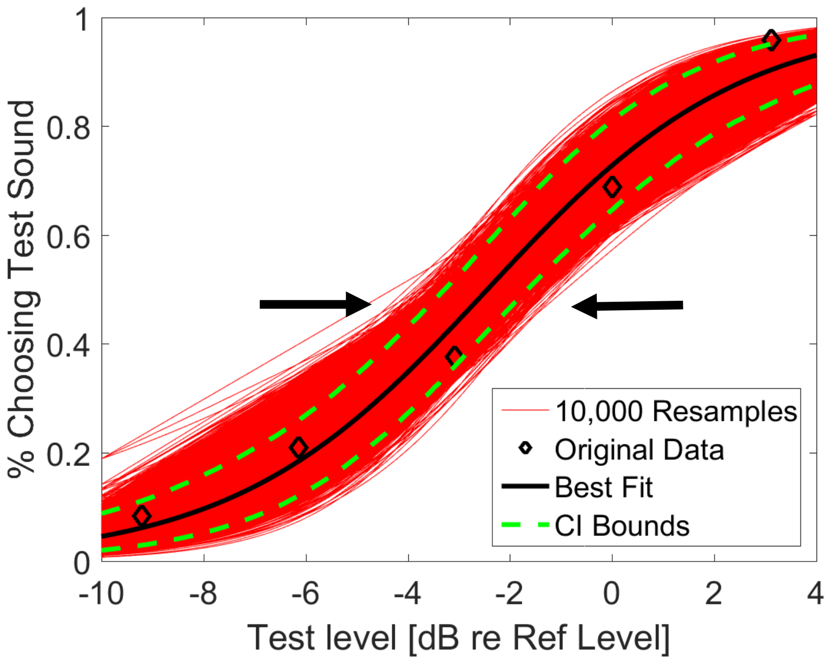


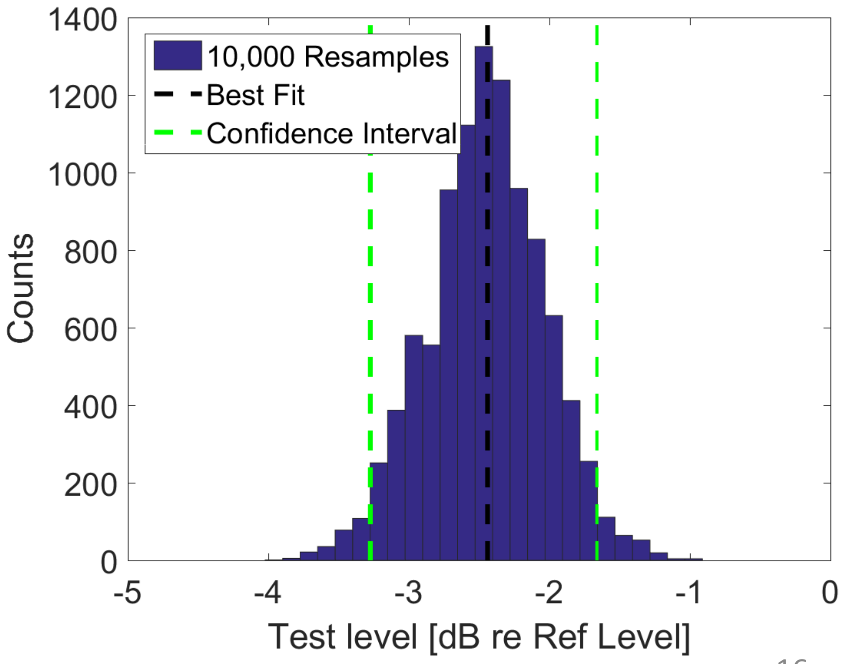
Bayesian Posterior Estimation

BPE numerically samples likelihood function…

allowing confidence intervals to be constructed

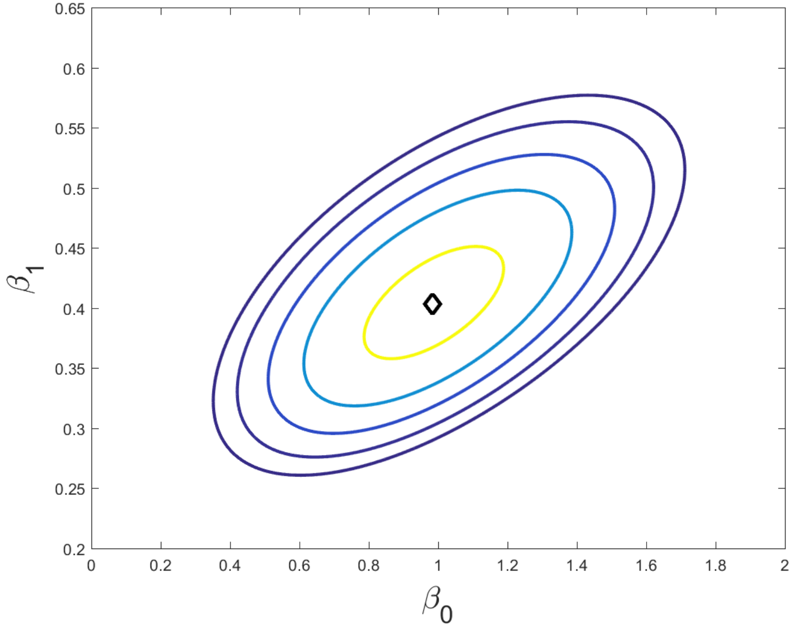
anywhere along logistic probability curve





Bayesian Posterior Estimation

All possible parameter combinations with corresponding goodness-of-fit yield "likelihood function"

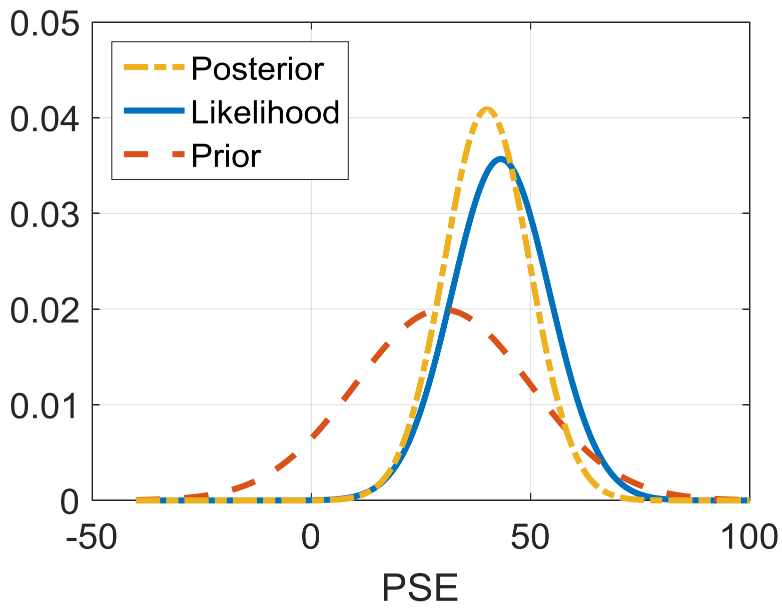


Bayesian Posterior Estimation

$$p(\beta_0, \beta_1 | Data) \propto L(Data | \beta_0, \beta_1) * p(\beta_0, \beta_1)$$

Posterior Likelihood Prior

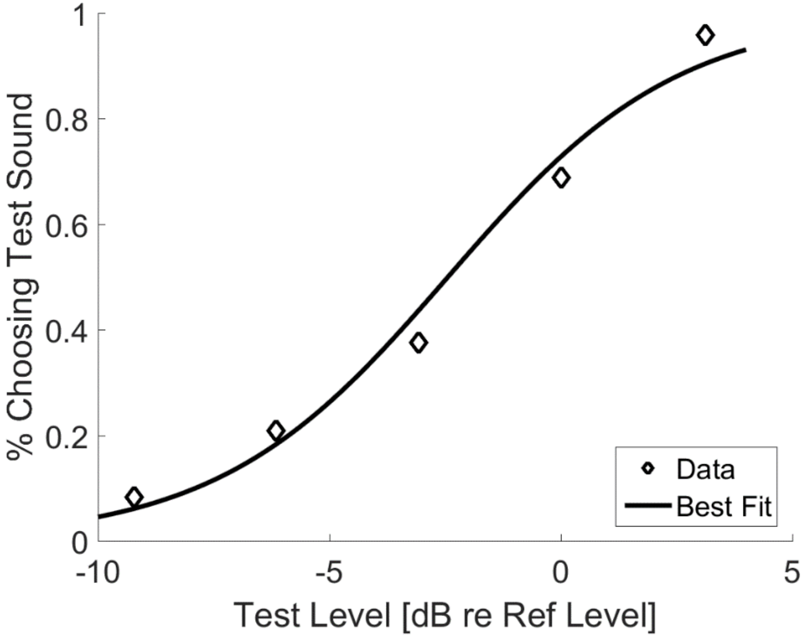
- BPE can include background knowledge (if known) in the form of “prior distributions”
- Previously posterior could only be evaluated when likelihood and prior known analytically
- MCMC methods enable numerical evaluation of arbitrary likelihoods/priors

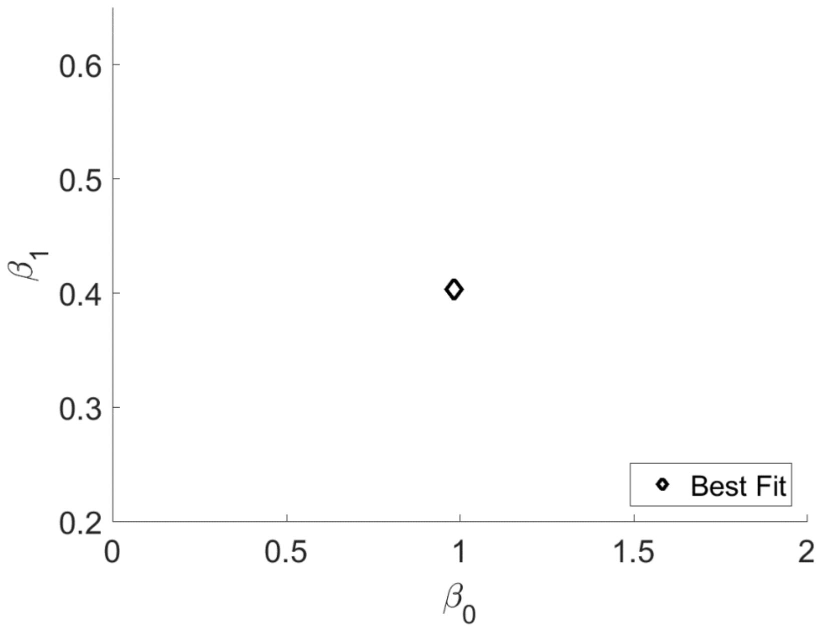


b. Bootstrap Analysis: Nonparametric

Bootstrap Analysis: Non-parametric

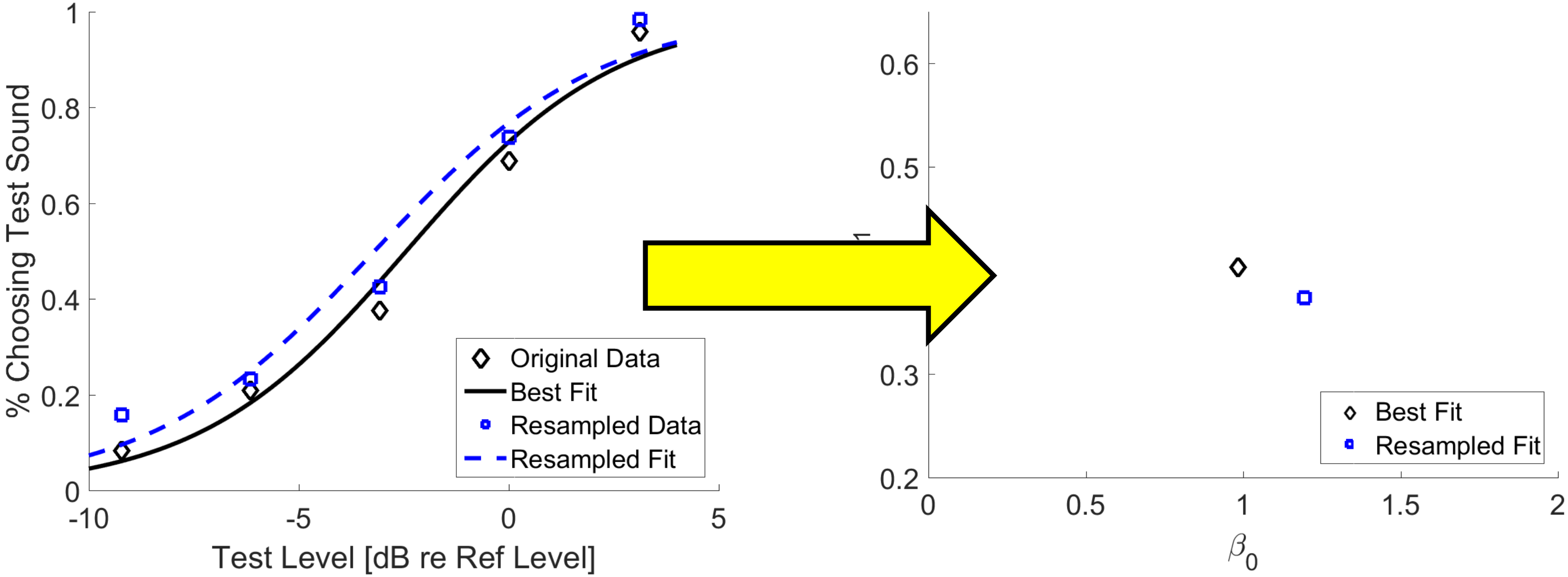
- What if we ran this experiment 10,000 times?



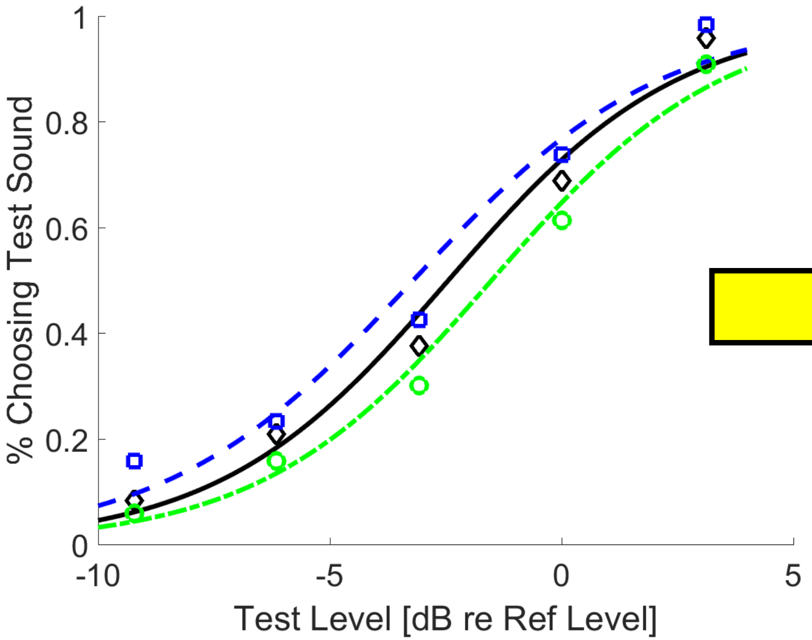


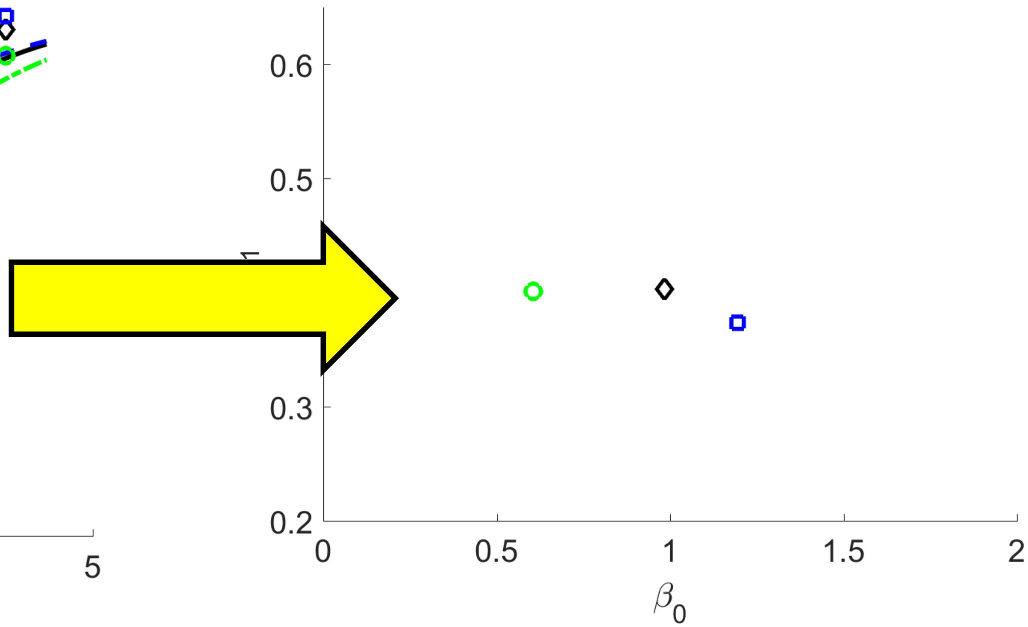
Bootstrap Analysis: Non-parametric

- Resample dataset with replacement
- Each resample uses *slightly less than* entire dataset

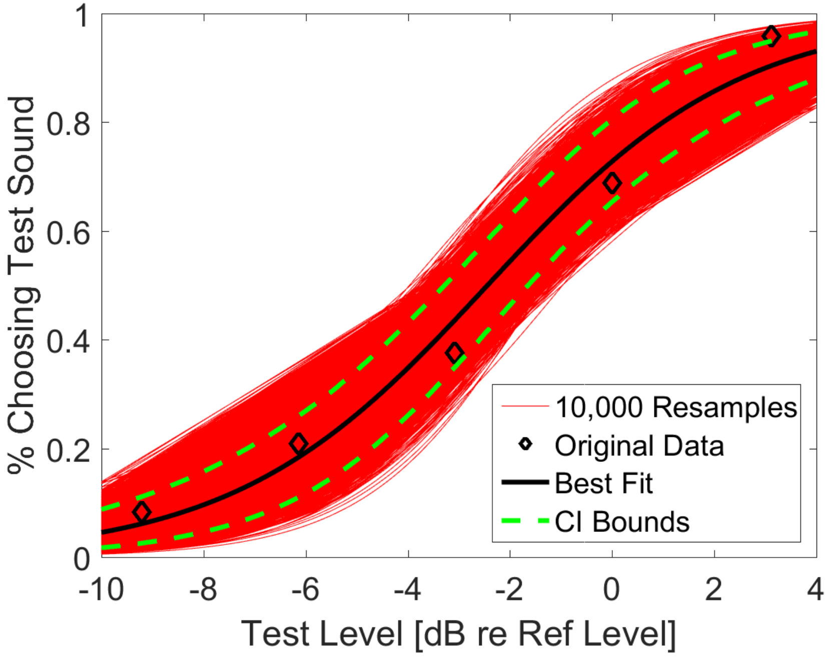


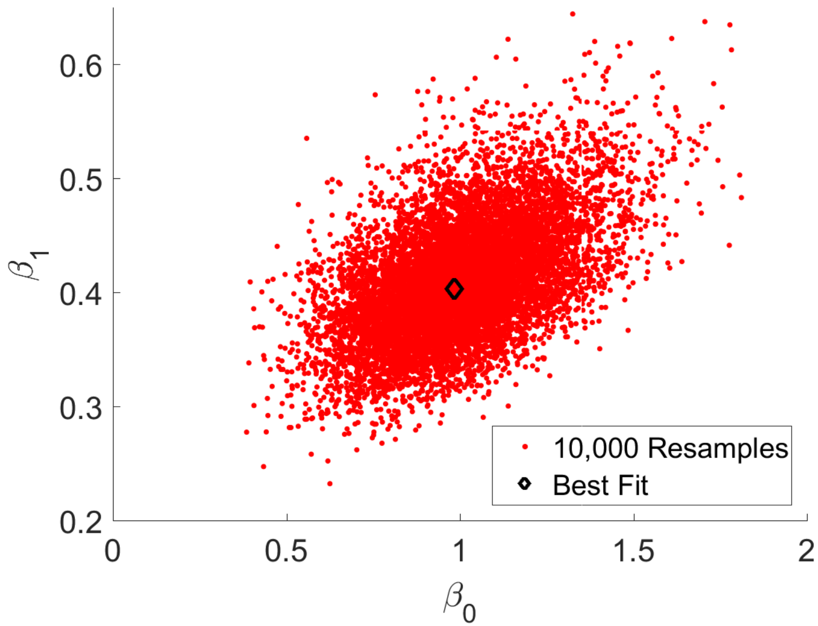
Bootstrap Analysis: Non-parametric



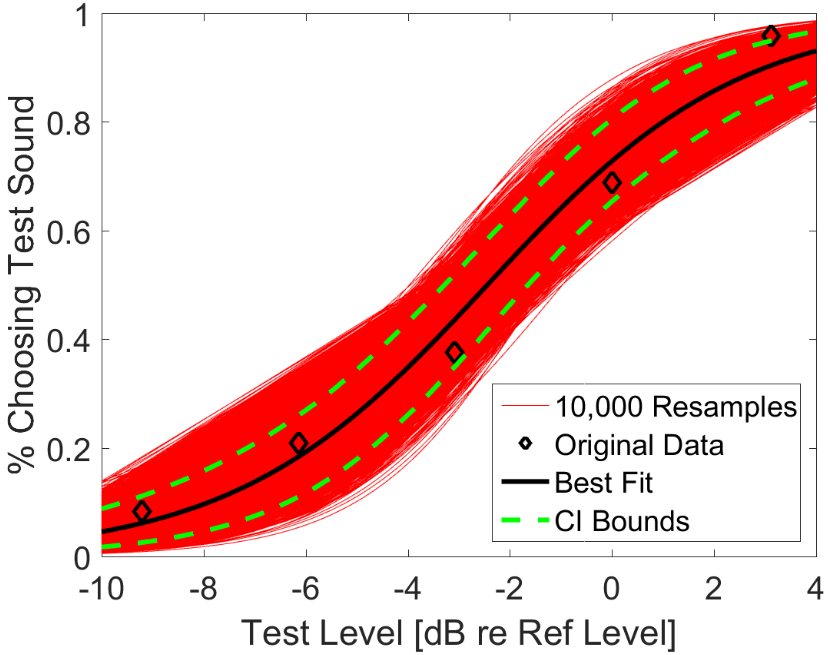


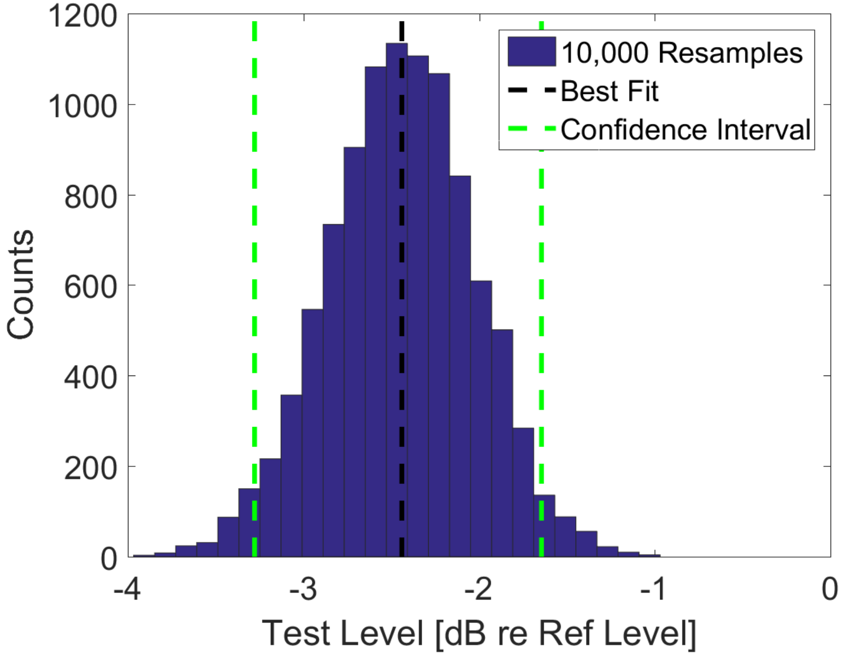
Bootstrap Analysis: Non-parametric





Bootstrap Analysis: Non-parametric

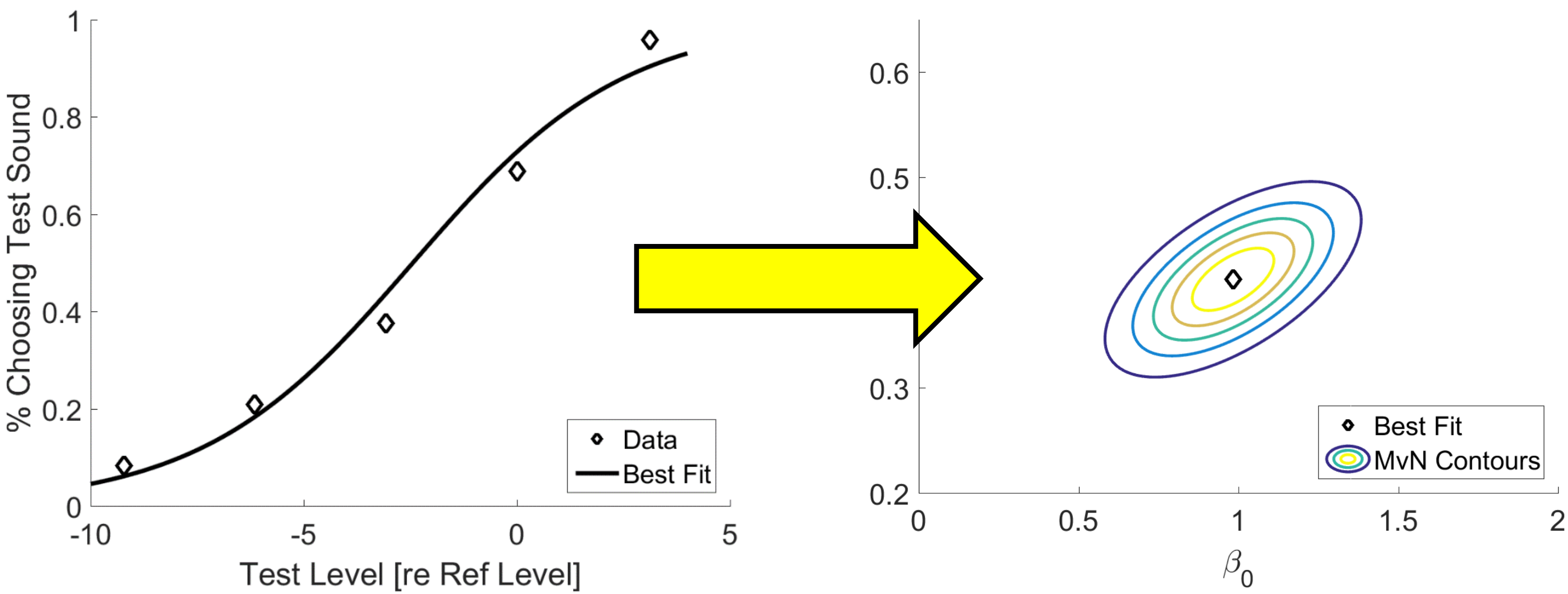




c. Bootstrap Analysis: Parametric

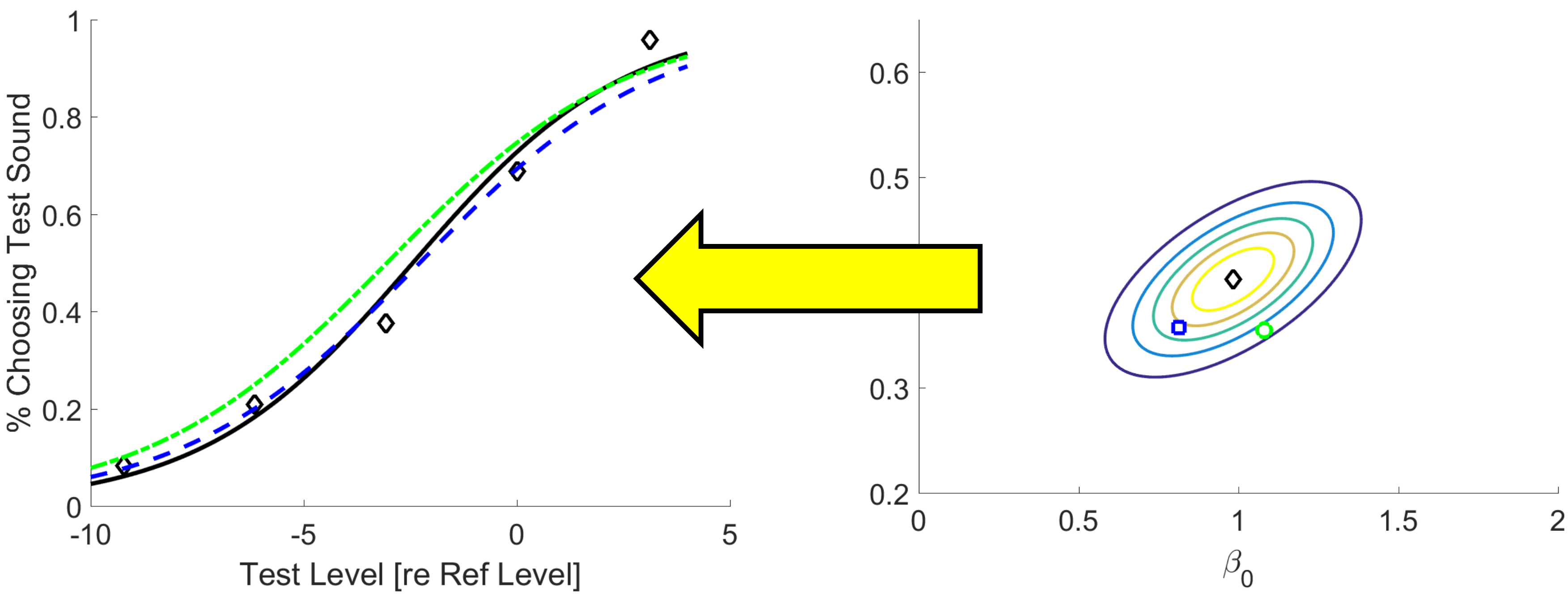
Bootstrap Analysis: Parametric

1) Fit data using maximum likelihood method (output is β_0, β_1, and $\mathrm{Cov}(\beta_0, \beta_1)$)
2) Use output to construct multivariate distribution

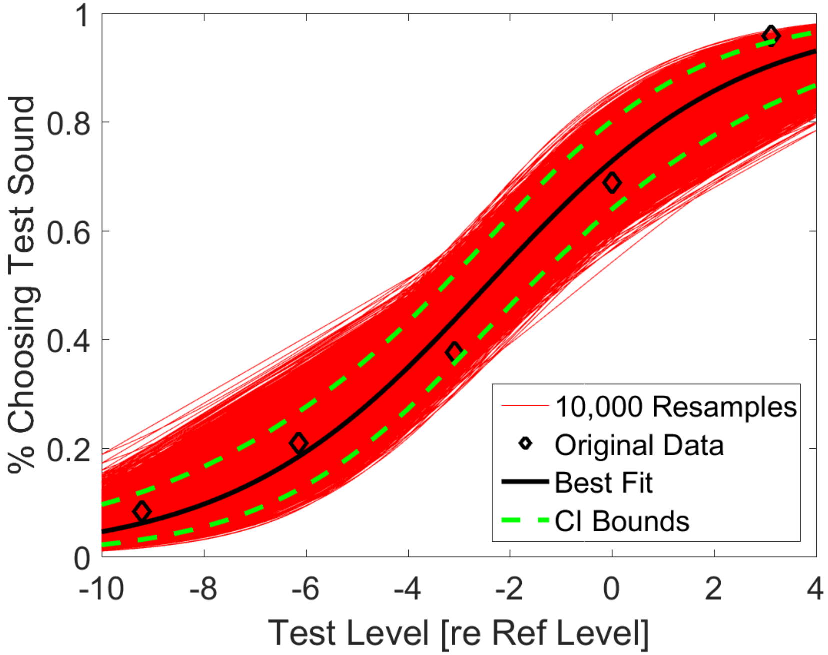


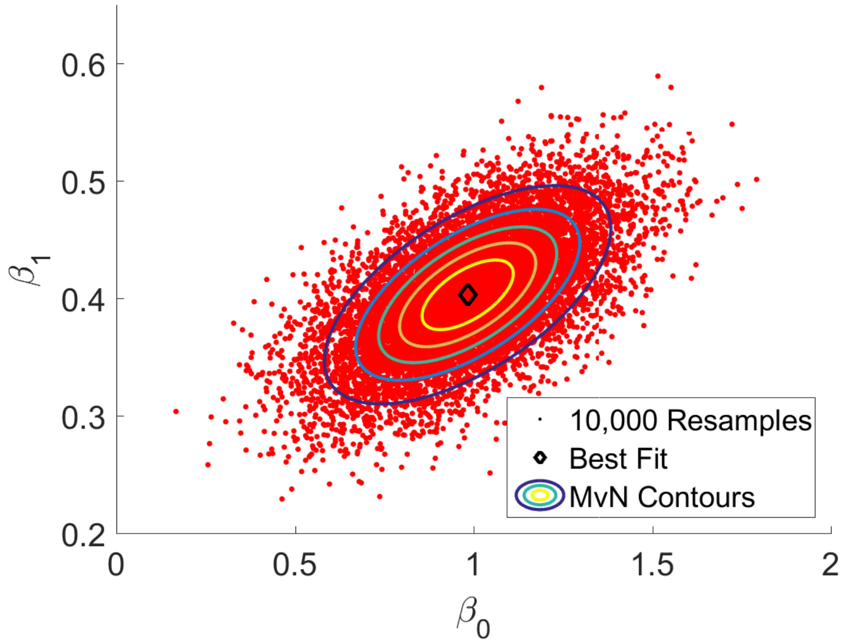
Bootstrap Analysis: Parametric

1) Fit data using maximum likelihood method (output is β_0, β_1, and $\mathrm{Cov}(\beta_0, \beta_1)$)
2) Use output to construct multivariate distribution
3) Sample from multivariate distribution

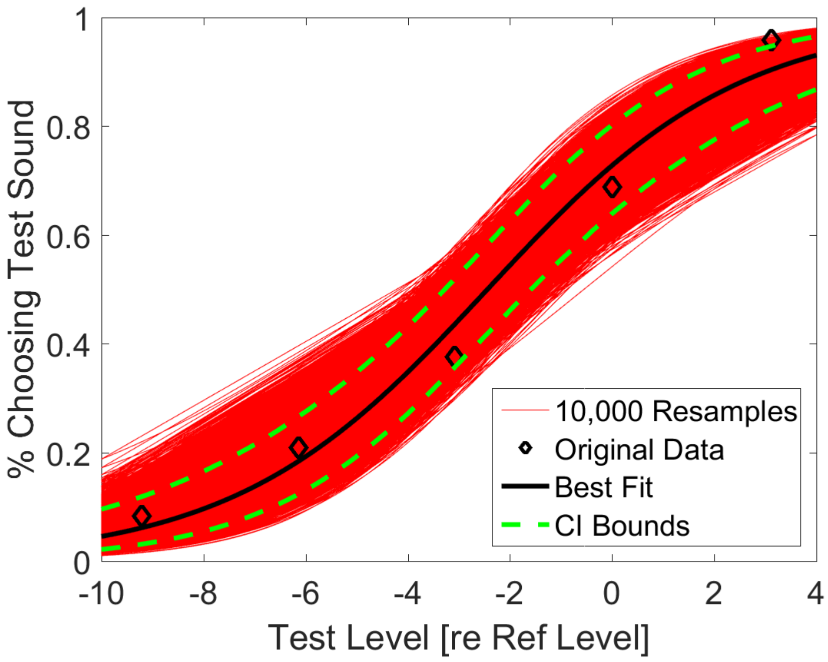


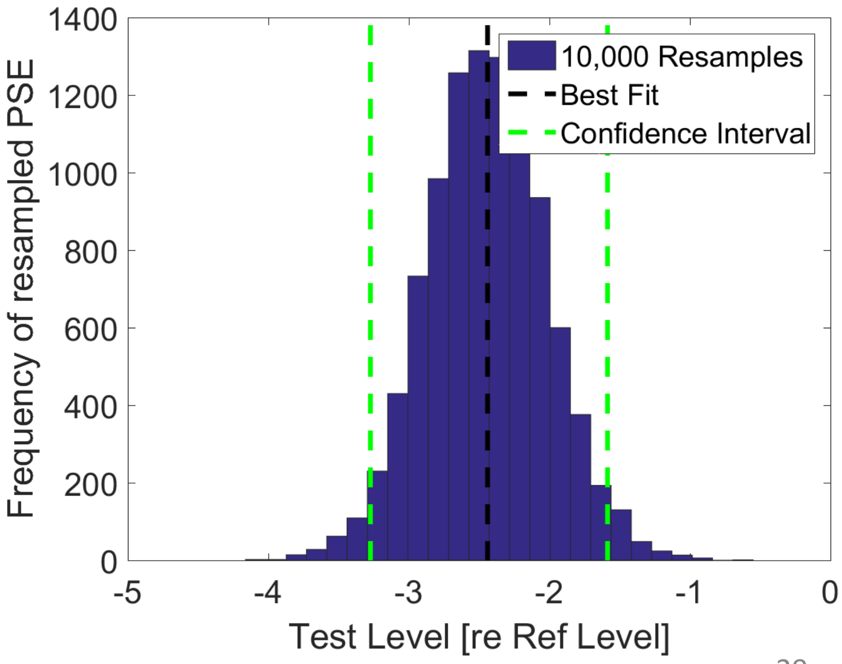
Bootstrap Analysis: Parametric





Bootstrap Analysis: Parametric





d. Delta Method

Delta Method: Theory

The GLM logistic regression model returns:

- β_0, β_1 -- maximum likelihood estimators of logistic regression parameters
- $\mathrm{Cov}(\beta_0, \beta_1)$ -- Covariance of parameters

Taylor Series Approximation to Variance of PSE [Morgan 1992]

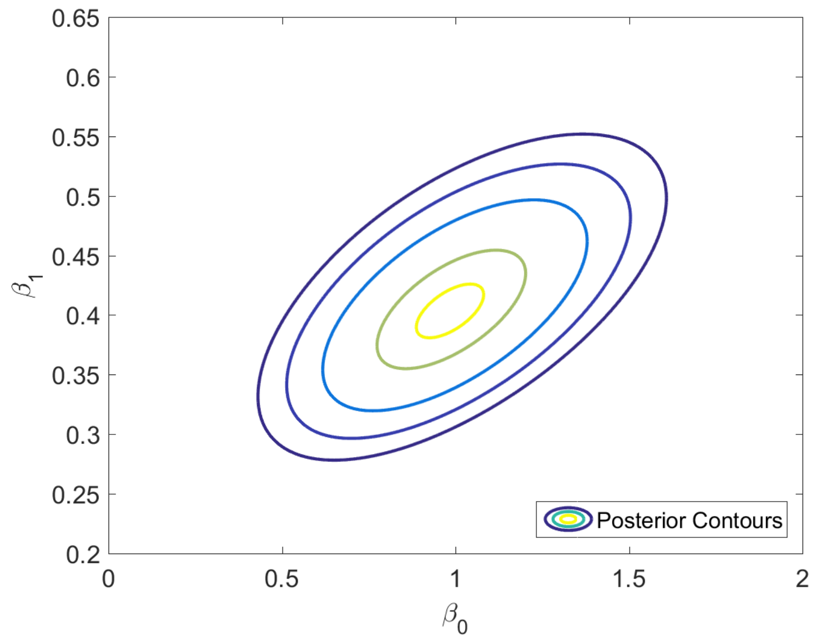
$$\mathrm{Var}\,(\mathrm{PSE}) = \frac{1}{\beta_1^2}\left[\mathrm{Var}(\beta_0) + \mathrm{PSE}^2 * \mathrm{Var}(\beta_1) + 2 * \mathrm{PSE} * \mathrm{Cov}(\beta_0, \beta_1)\right]$$

Delta Method Confidence Interval

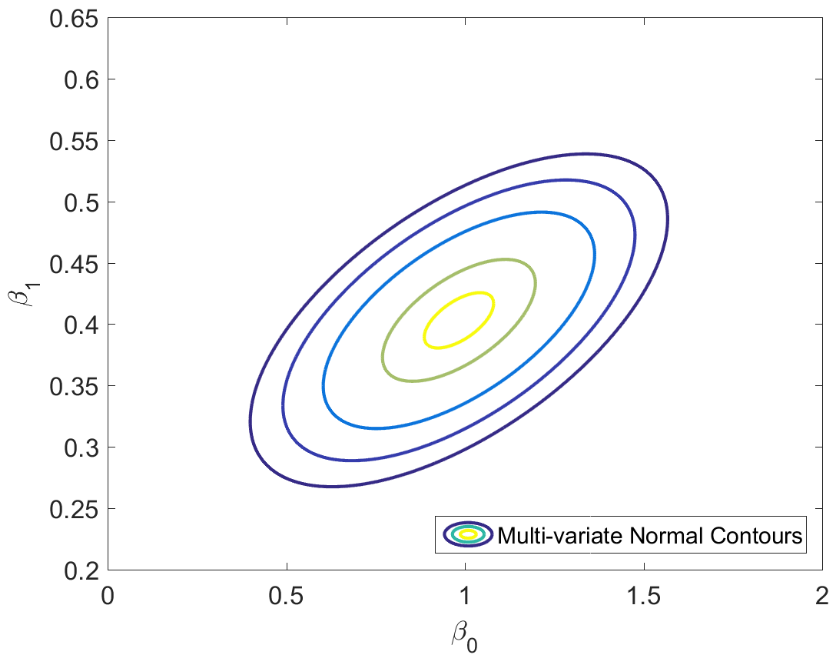
$$\mathrm{PSE} \pm z_{\left(1-\frac{\alpha}{2}\right)}\sqrt{\mathrm{Var}(PSE)}$$

Results

- All methods gave the same results!
  - PSE = -2.44 dB
  - 95% CI [-3.26, -1.62] dB



Bayesian Posterior Estimation



Parametric Bootstrap

Results: Guidance Table

Method	*Notes*
Bayesian Posterior Estimation	•Most flexible (can include prior information) •Uses all data for calculating likelihood •Diagnostics needed to ensure proper numeric performance
Bootstrap: Nonparametric	•Takes longest to calculate (10,000x as long as Delta Method) •Most affected by low-N binomial data
Bootstrap: Parametric	•Observable failure modes (e.g. negative slope)
Delta Method	•Closed form •Assumes confidence interval is symmetric about PSE •Unobservable failure modes

Conclusions

- Bayesian and Frequentist concepts yield same results

- What is most appropriate interval estimation technique among four standard solutions?

  -All methods yield equivalent results

  -Delta Method is fastest to calculate

  -BPE is most complex (pros and cons)

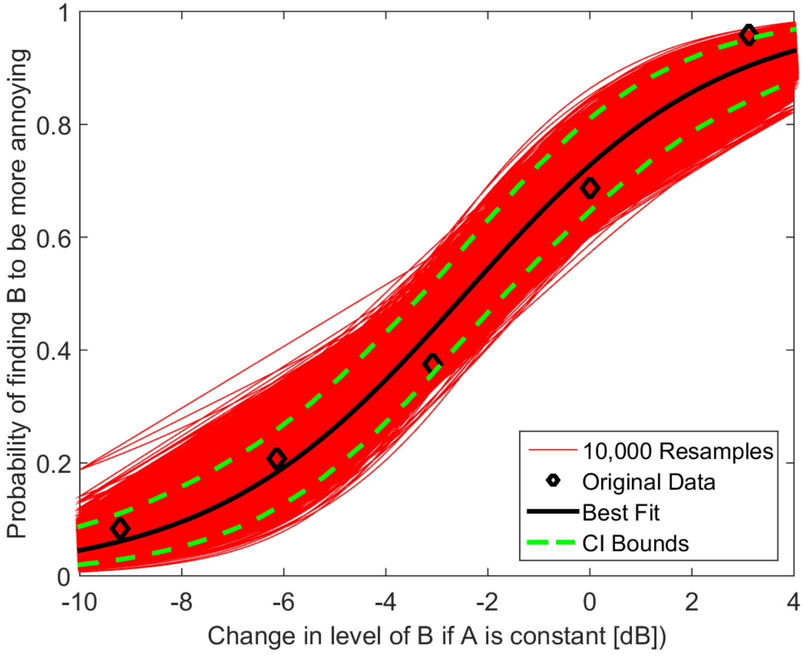
Thank You

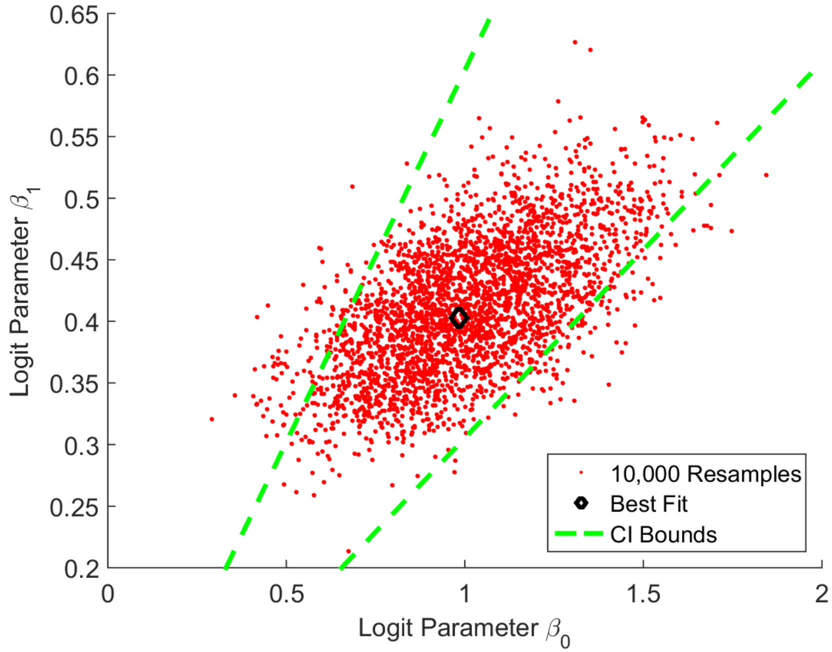
Reference:

- Morgan, B.J.T. Analysis of Quantal Response Data London: Chapman & Hall (1992).

Backup Slides

Bayesian Posterior Estimation





Results: Guidance Table

Method	*PSE*	*PSE Interval min—max*	*Longest Operation*	*Notes*
Delta	82.6	81.3—83.9	1 GLM fit (fastest)	•Closed form •Unknown failure modes
Bootstrap: Parametric	82.6	81.2—83.9	Sorting N resampled PSEs (2nd fastest)	•Resamples are normally distributed •Observable failure modes (e.g. negative slope)
Bootstrap: Nonparametric	82.6	81.3—83.9	N GLM fits (slowest)	•Fewest assumptions •Not suitable for low-N binomial data
Bayesian Posterior Estimation	82.6	81.4—83.9	N likelihood evaluations (2nd slowest)	•Most flexible (can include prior information) •Diagnostics needed to ensure proper MCMC performance

Bootstrap Analysis: Non-parametric